When do generalized entropies apply? How phase space volume determines entropy

Rudolf Hanel[1] and Stefan Thurner[1,2]
[1] *Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, A-1090, Austria*
[2] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*

We show how the dependence of phase space volume $\Omega(N)$ of a classical system on its size N uniquely determines its extensive entropy. We give a concise criterion when this entropy is not of Boltzmann-Gibbs type but has to assume a *generalized* (non-additive) form. We show that generalized entropies can only exist when the dynamically (statistically) relevant fraction of degrees of freedom in the system vanishes in the thermodynamic limit. These are systems where the bulk of the degrees of freedom is frozen and is practically statistically inactive. Systems governed by generalized entropies are therefore systems whose phase space volume effectively collapses to a lower-dimensional 'surface'. We explicitly illustrate the situation for binomial processes and argue that generalized entropies could be relevant for self organized critical systems such as sand piles, for spin systems which form meta-structures such as vortices, domains, instantons, etc., and for problems associated with anomalous diffusion.

PACS numbers: 05.20.-y, 02.50.Cw, 05.90.+m

Entropy relates the number of states of a system to an *extensive* quantity, which plays a fundamental role in its thermodynamical description. Extensive means that when two initially isolated systems A and B – with Ω_A and Ω_B the respective numbers of states – are brought in contact, the entropy of the combined system $A+B$ is $S(\Omega_{A+B}) = S(\Omega_A) + S(\Omega_B)$. Extensivity is not to be confused with *additivity* which is the property that $S(\Omega_A\Omega_B) = S(\Omega_A) + S(\Omega_B)$. Both, extensivity and additivity coincide if the number of states in the combined system is $\Omega_{A+B} = \Omega_A\Omega_B$. Clearly, for a non-interacting system Boltzmann-Gibbs (BG) entropy, $S_{\rm BG}[p] = \sum_i^\Omega g_{\rm BG}(p_i)$, with $g_{\rm BG}(x) = -x\ln x$, is simultaneously extensive *and* additive. By 'non-interacting' systems (short-range, ergodic, mixing, Markovian, ...) we mean $\Omega_{A+B} = \Omega_A\Omega_B$. For interacting statistical systems this is in general not true. If phase space is only partly visited this means $\Omega_{A+B} < \Omega_A\Omega_B$. In this case, it may happen that an additive entropy (such as BG) no longer is extensive and vice versa. With the hope to understand *interacting* statistical systems within a thermodynamical formalism and to ensure extensivity of entropy, so called *generalized entropies* have been introduced which usually assume trace form

$$S_{\rm gen}[p] = \sum_{i=1}^{\Omega} g(p_i) \quad , \qquad [\Omega \text{ ... number of states}] \quad (1)$$

where g is some function of p. It has been shown that g can not assume any functional form, but generalized entropies of trace form are restricted to the family of functions

$$S_{c,d}[p] \propto \sum_{i=1}^{\Omega} \Gamma(d+1, 1-c\log p_i) \quad , \qquad (2)$$

$\Gamma(.,.)$ being the incomplete gamma function, whenever a minimum set of requirements on g hold [1]. These requirements are the first three of the four Shannon-Khinchin (SK) axioms [2, 3], SK1: Entropy S depends continuously on p (g is continuous), SK2: entropy is maximal for the equi-distribution $p_i = 1/\Omega$ (g is concave. In physical systems this represents the equi-partition principle in micro-canonical ensembles), SK3: adding a zero-probability state to a system, $\Omega+1$ with $p_{\Omega+1} = 0$, does not change the entropy ($g(0) = 0$), and SK4: the entropy of a system – composed of sub-systems A and B – equals the entropy of A plus the expectation value of the entropy of B, conditional on A. If SK1-SK4 hold, the only possible entropy is BG [2, 3]. If only SK1-SK3 hold (additivity axiom violated) Eq. (2) is the generalized entropy with the constants (c,d) characterizing the universality class of entropy. $(c,d) = (1,1)$ is the class of BG entropy, $(c,d) = (q,0)$ is the class of Tsallis entropies. A universality class (c,d) not only characterizes the entropy of the system completely in the thermodynamic limit, it also specifies its distribution functions. Many recently introduced generalized entropic forms appear to be special cases of Eq. (2) [1]. The associated distribution functions are

$$\mathcal{E}_{c,d,r}(x) = e^{-\frac{d}{1-c}\left[W_k\left(B\left(1-\frac{x}{r}\right)^{\frac{1}{d}}\right) - W_k(B)\right]} \,, \qquad (3)$$

with $B \equiv \frac{(1-c)r}{1-(1-c)r}\exp\left(\frac{(1-c)r}{1-(1-c)r}\right)$, and as one possible choice, $r = (1-c+cd)^{-1}$, [1]. The function W_k is the k'th branch of the Lambert-W function, which is a solution of the equation $x = W(x)\exp(W(x))$. Only branch $k=0$ and branch $k=-1$ have real solutions W_k. Branch $k=0$ is necessary for all classes with $d \geq 0$, branch $k=-1$ for $d < 0$. The generalized logarithm for the entropy Eq. (2) is the inverse of $\mathcal{E} = \Lambda^{-1}$. Further properties of systems where SK1-SK3 hold are reported in [4].

It has often been argued that for statistical systems with strong and long-range correlations, Boltzmann-Gibbs statistical mechanics loses its applicability, and that under these circumstances generalized entropies become necessary. This is certainly not true in general. While correlations can be the reason for non-Boltzmann

distribution functions, BG entropy often remains the correct extensive entropy of the system [5].

In this paper we clarify the conditions under which BG entropy breaks down as the extensive entropy of a system. For ergodic systems, covering phase space, BG is always valid, regardless of what the correlations in the system might be. This was explicitly shown for binary systems in [5]. It is obvious that the *structure* of phase space, i.e. Gibbs Γ-space, is responsible for the Boltzmann-Gibbs framework to collapse and for generalized entropies to become necessary. Here we show that mere non-ergodicity is not enough: for generalized entropies to become necessary, Γ-space has to collapse in a specific way.

In the following we derive all results in terms of growth of phase space volume as a function of system size. We illustrate our results for binary systems where a graphical representation is possible in terms of decision trees. Binary systems with correlations [6, 7] have been studied in the light of generalized entropies in [5, 8–11]. On the basis of growth of Γ-space as a function of the number of states we present a set of concise criteria when generalized entropies are unavoidable and specify them by their universality classes.

What does extensivity mean? Consider a system with N elements, each of which can be in one of m states. The number of system configurations (microstates) are denoted by $\Omega(N)$, which depends on N in a system-specific way. Starting with Eq. (1) for equi-distribution, $p_i = 1/\Omega$ (for all i), we have $S_g = \sum_{i=1}^{\Omega} g(p_i) = \Omega g(1/\Omega)$. Extensivity for two subsystems A and B means that

$$\Omega_{A+B}\, g\left(1/\Omega_{A+B}\right) = \Omega_A\, g\left(1/\Omega_A\right) + \Omega_B\, g\left(1/\Omega_B\right) \quad . \quad (4)$$

Using the primary scaling property of generalized entropies $\lim_{x\to 0^+} \frac{g(\lambda x)}{g(x)} = \lambda^c$, (see [1]), we get asymptotically $g'(x) = cg(x)/x$, and Eq. (4) becomes

$$g'\left(1/\Omega_{A+B}\right) = g'\left(1/\Omega_A\right) + g'\left(1/\Omega_B\right) \quad . \quad (5)$$

The derivative of g is the generalized logarithm, $g'(x) = -\Lambda(x)$, and

$$\frac{1}{\Omega_{A+B}} = \mathcal{E}\left[\Lambda\left(\frac{1}{\Omega_A}\right) + \Lambda\left(\frac{1}{\Omega_B}\right)\right] = \frac{1}{\Omega_A} \otimes_g \frac{1}{\Omega_B} \quad . \quad (6)$$

A generalized product $\otimes_g$ can now be defined as $x \otimes_g y \equiv \mathcal{E}\left[\Lambda(x) + \Lambda(y)\right]$. If each 'particle' can be in one of m states, we finally get for the number of states in the system

$$\frac{1}{\Omega(N)} = \mathcal{E}\left[N\Lambda\left(\frac{1}{m}\right)\right] \quad , \quad (7)$$

or if we use the distribution functions and generalized logarithms of generalized entropies, Eq. (3), the number of microstates grows asymptotically as

$$\Omega(N) = \frac{1}{\mathcal{E}_{c,d}(\mu(c-1)N)} = \exp\left[\frac{d}{1-c} W_k\left(\mu(1-c)N^{\frac{1}{d}}\right)\right] \quad (8)$$

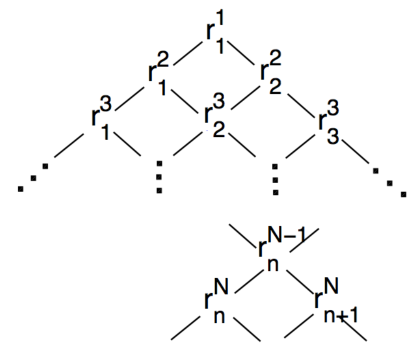

FIG. 1: Decision tree (triangle) for binary processes representing the probabilities r_n^N of n heads and $N-n$ tails after N throws. r_n^N is the probability of a *specific* sequence, $\{\varphi_1, \varphi_2, \ldots, \varphi_N\}$, which does not depend on the order of events, but only on the number of n events in state $\varphi_i = 1$ and $N-n$ events in state $\varphi_j = 0$. Leibnitz rule (scale invariance) [8] holds if $r_n^N + r_{n+1}^N = r_n^{N-1}$ at all levels N and $N-1$.

where μ is some positive constant. At this stage note that for all non-BG systems, i.e. $(c,d) \neq (1,1)$, the number of states $\Omega(N)$ grows sub-exponentially with N.

Inversely, given the phase space volume as a function of system size, we can now compute the generalized entropy, i.e. its universality class characterized by (c,d). Again using the primary scaling property for generalized entropies and de L'Hospital rule $\lambda^c = \lim_{x\to 0^+} \frac{g(\lambda x)}{g(x)} = \frac{\lambda g'(\lambda x)}{g'(x)} = \frac{\lambda \Lambda(\lambda x)}{\Lambda(x)}$, together with Eq. (7) we get for large $\Omega(N)$

$$\lambda^{c-1}\Lambda\left(1/\Omega(N)\right) = \Lambda\left(\lambda/\Omega(N)\right) \quad , \quad (9)$$

or $\lambda^{\frac{1}{c-1}} = \lim_{N\to\infty} \frac{\Omega(\lambda N)}{\Omega(N)}$, which can be simplified to

$$\frac{1}{1-c} = \lim_{N\to\infty} N \frac{\Omega'(N)}{\Omega(N)} \quad . \quad (10)$$

For d we use the secondary scaling relation for generalized entropies [1], $(1+a)^d = \lim_{x\to 0} \frac{g(x^{1+a})}{x^{ac} g(x)}$. Taking the derivative with respect to a on both sides we get

$$d(1+a)^{d-1} = (1+a)^d \lim_{x\to 0} \log x \left(\frac{x^{1+a} g'(x^{1+a})}{g(x^{1+a})} - c\right) . \quad (11)$$

Set $a \to 0$, and use de L'Hospital rule to get $d = \lim_{x\to 0}(1 - c + xg''(x)/g'(x))$ so that with Eq. (7) we have

$$d = \lim_{N\to\infty} \log \Omega \left(\frac{1}{N}\frac{\Omega}{\Omega'} + c - 1\right) \quad . \quad (12)$$

To see how phase space collapses for generalized entropy systems we illustrate the above result in the context of binary sequences, e.g. $\varphi = \{0,1,1,0,0,1,\cdots\}$. Let $\varphi|_N = \{\varphi_i\}_{i=1}^N$ denote a sequence of length N.

Any correlations in sequences (on all levels of N) [5] are completely determined by the set of joint probability functions $\{p_N\}_{N=1}^{\infty}$, where $p_N(\varphi_1, \varphi_2, \ldots, \varphi_N)$ is the joint probability of a sequence of length N. A sequence $\varphi|_N$ contains $k(N) = \sum_{i=1}^{N} \varphi_i$ 'ones' and $N - k(N)$ 'zeros'. If $p_N(\varphi_1, \varphi_2, \ldots, \varphi_N)$ is totally symmetric in its arguments the probability of sequences of length N depends on k only, $r_k^N = p_N(\varphi_1, \varphi_2, \ldots, \varphi_N)$. There exist $\binom{N}{k}$ sequences $\varphi|_N$ with k 'ones' and $N-k$ 'zeros' and $\sum_{k=0}^{N} \binom{N}{k} r_k^N = 1$. The r_k^N can be arranged into a triangle, Fig. 1, representing the equi-probable sequences of binary events in a 'decision tree'. The full phase space volume (on level N) is

$$\Gamma_N = \{0,1\}^N \qquad , \qquad \Omega(N) = |\Gamma_N| = 2^N \quad . \tag{13}$$

In other words, the number of microstates (number of sequences up to level N) is $\Omega(N) = 2^N$. Obviously the number of states increases exponentially and when substituted in Eqs. (10) and (12) we recover $(c,d) = (1,1)$, i.e. BG entropy.

We now introduce restrictions to phase space such that not all sequences are allowed on all levels N anymore. We denote the number of states in the restricted phase space by $\Omega^{(R)}(N) = |\Gamma_N^{(R)}|$, and can immediately discuss an interesting fact. Imagine that phase space of sequences is extremely confined, say to a situation where in the thermodynamic limit all sequences approach a common point (same number of 'ones' in the sequence) $\xi = \lim_{N\to\infty} k(N)/N$. In all cases where this point ξ is neither 0 or 1, BG is the only possible extensive entropy. This can be formulated in a

Theorem: Define a restricted phase space $\Gamma_N^{(R)}$ on level N by

$$\Gamma_N^{(R)} \equiv \left\{ \varphi|_N \in \Gamma_N \ : \ \underline{K}(N) \leq \sum_{i=1}^{N} \varphi_i \leq \overline{K}(N) \right\} \quad ,$$

i.e., at level N the system only allows sequences $\varphi|_N$ with more than $\underline{K}(N)$ and less than $\overline{K}(N)$ 'ones'. If the restriction of phase space is such that $\lim_{N\to\infty} \frac{\underline{K}(N)}{N} = \lim_{N\to\infty} \frac{\overline{K}(N)}{N} = \xi$, where $\xi \in (0,1)$, then asymptotically the number of states grows exponentially ($\Omega(N) = b^N$ for some $b > 0$), and the extensive entropy is BG.

The proof is to show that both, lower and upper bounds for $\Omega(N)$, yield BG. The theorem states that for generalized entropies to exist it is necessary that the sequences are constrained to the situation where either $\lim_{N\to\infty} \frac{\overline{K}(N)}{N} = 0$, or $\lim_{N\to\infty} \frac{\underline{K}(N)}{N} = 1$. In other words the sequences are asymptotically confined to a region of measure zero around the flanks of the decision triangle, i.e. the *boundary* of phase space. The theorem has two further implications:

1. In case of probability distributions p_N which are *not* totally symmetric in their arguments, generalized entropies can exist even though phase space need not be limited to the boundary of the decision triangle (as in the theorem). If the number of sequences $\varphi \in \Gamma^{(R)}$ (i.e. the number of free decisions) up to level N grows sufficiently sub-linearly with N, then the limit-points of sequences may be found along the entire base of the decision triangle (compare remark on super-diffusive random walks below). This means that the multiplicity of sequences with k out of N 'ones' in the large N limit grows sufficiently slower than the binomial multiplicity for totally symmetric p_N.

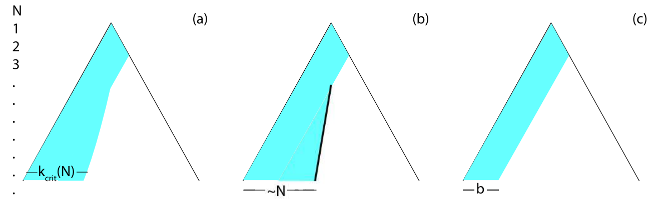

FIG. 2: Binary decision trees: (a) Schematic view of allowed sequence regions. If sequences are confined to the shaded regions (left of critical sequence line k^{crit}) the extensive entropy of the system is not BG but a generalized entropy. (b) Maximum line k^{crit} at which BG entropy starts. Any system containing this line or sequences to the right of it, will be BG. (c) If the region is confined to a strip of size b, the extensive entropy is Tsallis entropy, $S_{q,0}$, with $q = 1 - \frac{1}{b}$.

2. The theorem can trivially be generalized from binary processes to m-state systems. This is done by passing from the binomial to a multinomial description.

We now show how different restrictions on phase space lead to various specific generalized entropies. We assume the existence of a critical sequence φ^{crit} which follows the path $k^{\mathrm{crit}}(N)$, see Fig. 2 a. This means that after N steps the sequence has produced a maximum of $k^{\mathrm{crit}}(N)$ 'ones'. To the right of this sequence all sequences are forbidden. The phase space volume of such systems grows like [15]

$$\Omega^{(R)}(N) = \sum_{i=1}^{k^{\mathrm{crit}}(N)} \binom{N}{i} \quad . \tag{14}$$

For any k, $\lim_{N\to\infty} \sum_{i=1}^{k} \binom{N}{i} / \binom{N}{k} = 1$, which allows to asymptotically approximate Eq. (14)

$$\Omega^{(R)}(N) \approx \binom{N}{k^{\mathrm{crit}}(N)} \quad . \tag{15}$$

Using Stirling's formula, taking logs on both sides and keeping terms to leading order we arrive at

$$k^{\mathrm{crit}}(N) \approx N \exp\left[W_{-1}\left(-\frac{1}{N} \log \Omega(N) \right) \right] \quad . \tag{16}$$

This means that for any system whose sequences are confined to regions left to the critical sequence $k^{\mathrm{crit}}(N)$, generalized entropies as specified in Eq. (8) are necessary. We now discuss some examples.

Maximum restricted phase space. Consider $(c,d)=(1,1)$, i.e. $\Omega(N)=2^N$. From Eq. (16) we get $k^{\rm crit}(N)\approx N$. This means that for systems with generalized entropies $k^{\rm crit}(N)$ grows in a sufficiently sub-linear way with N, e.g. $k^{\rm crit}(N)\propto N^\alpha$ with $0<\alpha<1$. If $\alpha=1$ and $k^{\rm crit}(N)=\varepsilon N$, no matter how small $\varepsilon>0$, the system belongs to BG.

Power-law growth. For a power-like growth of phase space, $\Omega(N)=N^b$, we have $k^{\rm crit}(N)\approx N\exp[W_{-1}(\frac{b}{N}\log\frac{b}{N}-\frac{b}{N}\log b)]$. Expanding the Lambert-W function we get $k^{\rm crit}(N)\approx b\exp(-\frac{\log b}{1-\log\frac{b}{N}})\to b$, in the large N limit. The phase space collapse is seen in the decision triangle as a restriction to a strip of width b, Fig. 2 c. In this case $(c,d)=(1-\frac{1}{b},0)$, i.e. Tsallis entropy applies exactly. This is a well known result [8, 12].

Streched exponential growth. For stretched exponential growth $\Omega(N)=\exp(\lambda N^\gamma)$, Eq. (16) can be rewritten to $k^{\rm crit}(N)\approx\log\Omega/W_{-1}[-\log\Omega/N]$ and the Lambert-W term is reasonably approximated by $\log(N/\log\Omega)+\log(\log(N/\log\Omega))$. With this $k^{\rm crit}(N)\approx\frac{\lambda}{1-\gamma}\frac{N^\gamma}{\log(N)}$, and the entropy is $(c,d)=(1,1/\gamma)$.

Note that systems with confined areas in their decision trees are examples for strong memory. The system has to remember how many 'ones' have occurred in its trajectories, see examples in [9, 10]. Inversely, given a critical sequence line $k^{\rm crit}(N)$, the universality class of the corresponding generalized entropy (c,d) can be computed.

Given the dependence of phase space volume Ω on system size we showed how to determine the associated extensive generalized entropy by computing the exponents (c,d). We demonstrated that different generalized (non-additive) entropies – i.e. $(c,d)\neq(1,1)$ – correspond to different ways of sub-exponential growth of Γ-space. We related the growth of phase space volume to the increase of the number of *statistically relevant* (dynamical) micro-states in the system. We found that whenever the fraction of dynamical variables $\frac{k}{N}$ vanishes for large N, $\lim_{N\to\infty}\frac{k}{N}=0$, generalized entropies become unavoidable. This extreme confinement of relevant variables to a set of measure zero means that almost all states in the system are the same, or equivalently, the bulk of the degrees of freedom is frozen. In other words, statistically relevant activity happens within a tiny fraction, $\frac{k}{N}$, of the system which can be seen as a collapse of phase space volume to some low-dimensional 'surface'.

In conclusion we hypothesize that generalized entropies are relevant for physical systems being dominated by 'surface effects', including the following:

• *Self organized critical systems.* In sandpiles consider discrete sites where sand grains can be. The (binary) state of a site is being occupied by a grain or not. In a sandpile the bulk of the system is occupied and just the surface of the pile contains its statistically relevant degrees of freedom. The trajectory of a sand grain in a classical sandpile model follows sequences much alike those in the decision tree, Fig. 2 c.

• *Spin systems with dense meta-structures*, such as spin-domains, vortices, instantons, caging, etc. If these meta-structures bind a vast majority of spins into (metastable) objects, the remaining spins – not belonging to these structures – can move freely only in surface-like regions between these objects. For instance spin systems on random networks growing with constant connectedness (number of links divided by number of nodes squared) can be shown to require Tsallis entropy.

• Super-diffusion. Consider a one dimensional accelerating random walk, where each left-right decision is followed by N^β $(0<\beta<1)$ steps in that direction (N being the total number of steps the walk has so far taken). This process is a super-diffusive process ($\langle x^2(t)\rangle\propto t^{2-\beta}$) which requires a generalized entropy of type $(c,d)=(1,1/\beta)$.

• *Anomalous diffusion.* The presented results could also apply whenever states of a statistical system are excluded by the presence of other materials restricting mobility in Euclidean space. Think e.g. of diffusion in porous media where statistically relevant action takes place on restricted surface-like areas, and not in full 3D.

For non-commutative variables alternative routes to generalized entropies may exist [13, 14].

This work was inspired by Prof. C. Tsallis to whom we are deeply indebted for countless discussions, suggestions and his fantastic hospitality at CBPF. We acknowledge support by Faperj and CNPq (Brazilian agencies).

[1] R. Hanel, S.Thurner, *Europhys. Lett.* **93**, 20006 (2011).
[2] C.E. Shannon, *The Bell System Technical Journal* **27**, 379,623 (1948).
[3] A.I. Khinchin, *Mathematical foundations of information theory* (Dover Publ., New York), (1957).
[4] R. Hanel, S. Thurner, M. Gell-Mann, *Proc. Nat. Acad. Sci.* doi/10.1073/pnas.1103539108 (2011).
[5] R. Hanel, S.Thurner, C. Tsallis, *Eur. Phys. J.* B **72**, 263-268 (2009).
[6] B. de Finetti, Atti della R. Academia Nazionale dei Lincei 6, Memorie, Classe di Scienze Fisiche, Mathematice e Naturale, **4** 251-299 (1931).
[7] Jaynes E.T., in *Bayesian Inference and Decision Techniques with Applications*, Goel P.K. and Zellner A., eds. (North-Holland, Amsterdam, 1986), p. 31.
[8] C. Tsallis , M. Gell-Mann, Y. Sato, *Proc. Nat. Acad. Sci.* **102**, 15377 (2005).
[9] J.A. Marsh, M.A. Fuentes, L.G. Moyano, C. Tsallis, *Physica* A **372**, 183-202 (2006).
[10] L.G. Moyano, C. Tsallis, M. Gell-Mann, *Europhys. Lett.* **73**, 813 (2006).
[11] A. Rodriguez, V. Schwammle, C. Tsallis, *JSTAT*, P09006

(2008).

[12] Tsallis C., in *Nonextensive Entropy* M. Gell-Mann, C. Tsallis, eds., (Oxford University Press), (2004).

[13] F. Caruso, C. Tsallis, *Phys. Rev.* E **78**, 021102 (2008).

[14] A. Saguia, M.S. Sarandy, *Phys. Lett. A* **374**, 3384-3388 (2010).

[15] For m states the critical number is obtained similarly: $\Omega^{(R)}(N) = \sum_{j=1}^{k^{\mathrm{crit}}(N)} \sum_{|n|=j} \binom{N}{n*}$, with the multi-indices $n = (n_1, n_2 \ldots, n_{m-1})$ and $n^* = (n_1, n2, \ldots, n_{m-1}, N - |n|)$, with $|n| = \sum_{i=1}^{m-1} n_i$. The multinomial factor is $\binom{N}{n^*} = N!/\prod_{j=1}^{m} n_j^*!$.